HIDING DATA IN VoIP

Józef Lubacz, Wojciech Mazurczyk, and Krzysztof Szczypiorski*
Warsaw University of Technology, Institute of Telecommunications
Warsaw, Poland, 00-665, Nowowiejska 15/19

ABSTRACT

The paper presents and compares steganographic techniques that can be used to enable hidden communication within computer networks. A new class – "steganophony" – of such methods is introduced. Specific methods proposed by the authors – LACK, HICCUPS and SIP-based VoIP protocols' steganography – are briefly described.

1. INTRODUCTION

The main aim of network steganography is to hide secret data in users' normal data transmissions, ideally, so it cannot be detected by third parties. One of the most popular steganographic techniques is to use a covert channel, which enables manipulating certain properties of the communications medium in an unexpected, unconventional, or unforeseen way. In the past few years the interest in steganographic methods that may be used in computer networks has grown considerable, mostly due to presumed usage of hidden communication by terrorists. Many new methods have been proposed and analyzed (Zander and Armitage, 2007, Petitcolas et al., 1999, Murdoch et al., 2005), also in meaning of translation among heterogenous environments (Szczypiorski et al., 2007, Szczypiorski et al., 2008). In this paper we propose a classification of network steganography methods.

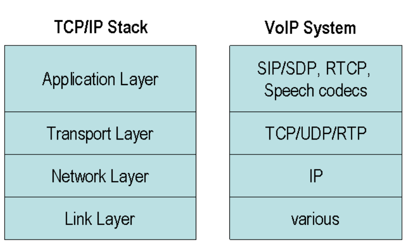

Fig. 1 VoIP stack and protocols

Voice over IP (VoIP), or IP telephony, is one of the services of the IP world which is changing the entire telecommunications landscape. Because of its popularity, it is becoming a natural target for steganography. We propose to name steganographic techniques applied to VoIP traffic **steganophony**. This term pertaining to information-hiding techniques in any layer of the TCP/IP protocol-stack (Fig. 1), including also methods such as audio watermarking and techniques applied in speech codecs.

For VoIP systems, four possible hidden communication scenarios may be considered, as illustrated in Fig. 2. The first scenario (marked with 1 in Fig. 2) is most common: the sender and the receiver perform VoIP conversation while simultaneously exchanging steganograms. The conversation path is the same as the hidden path. For the next three scenarios (marked 2-4 in Fig. 2) only a part of the VoIP end-to-end path is used for hidden communication as a result of actions undertaken by intermediate nodes; the sender and receiver are, in principle, unaware of the steganographic data exchange

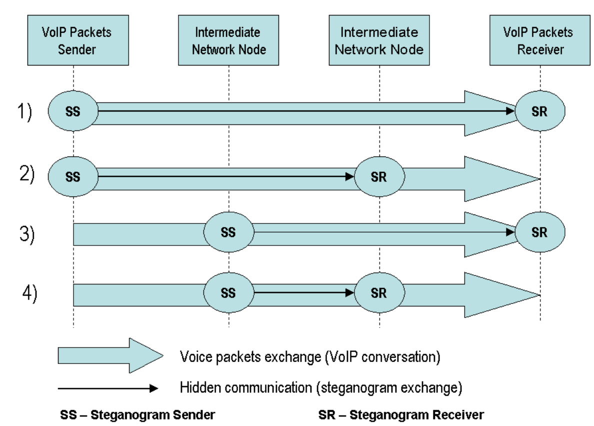

Fig. 2 Hidden communication scenarios for VoIP

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **01 DEC 2008**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**

4. TITLE AND SUBTITLE **Hiding Data In VoIP**	5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Warsaw University of Technology, Institute of Telecommunications Warsaw, Poland, 00-665, Nowowiejska 15/19**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM002187. Proceedings of the Army Science Conference (26th) Held in Orlando, Florida on 1-4 December 2008, The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**6**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

2. CLASSIFICATION OF NEWTORK STEGANOGRAPHY

Generally, network steganography may be divided into two broad groups: (S1) steganographic methods that modify packets and (S2) methods that modify packets' time relations (Fig. 3).

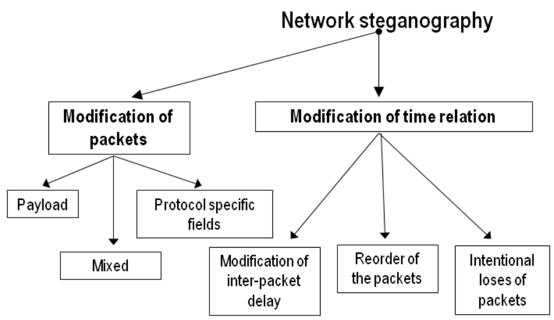

Fig. 3 Network steganography classification

Steganophony (i.e. VoIP steganography) techniques can be classified into three groups (Fig. 4):

(S1) Steganographic methods which modify packets – network protocol headers or payload fields. Examples of such solutions include (1) modifications of free/redundant headers' fields of IP, UDP or RTP (Schulzrinne et al., 2003) protocols during conversation phase and (2) modification of signaling messages in e.g. SIP (Rosenberg et al., 2002). Information hiding which is based on affecting packets' payload usually uses digital audio watermarking algorithms, e.g. DSSS (Cox et al. 1997) and QIM (Chen and Wornell, 2001).

(S2) Steganographic methods which modify packets' time relations, e.g. by affecting sequence order of RTP packets (Kundur and Ahsan, 2003), modifying their inter-packet delay (Berk et al., 2005) or by introducing intentional losses (Servetto et al., 2001).

(S3) Hybrid steganographic methods which modify both the content of packets and their time relations. An example of such solution is the LACK (Lost Audio Packets Steganography) method, which is described in the third section of this paper.

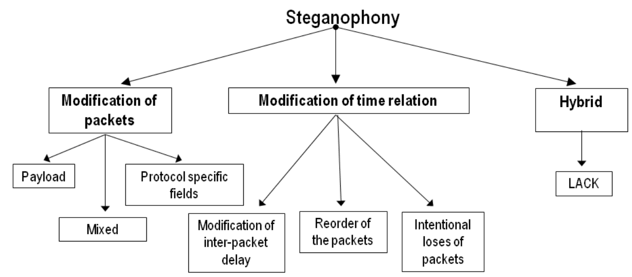

Fig. 4 Steganophony classification

Examples of the steganographic methods from group S1:

- Methods which modify protocols specific fields – SIP, SDP (Handley, 2006), RTP, RTCP (VoIP specific protocols) and additionally: IP, TCP, UDP (network specific protocols).
- Methods which modify packet's payload: audio watermarking algorithms, speech codec steganographic techniques (e.g. using SID frames or codec specific steganographic methods).
- Mixed techniques: HICCUPS (Hidden Communication System for Corrupted Networks, Szczypiorski, 2003).

Some characteristic features of the above methods:

- Steganographic methods which use protocol specific fields usually yield relatively high steganographic capacity. Implementation and detection is relatively straightforward. Drawback: potential loss of some of the protocols' functionality.
- Steganographic methods which utilize payload of the packets generally yield lower steganographic capacity and are harder to implement and detect. Drawback: potential deterioration of voice quality.
- Mixed techniques offer high steganographic capacity, but the implementation is harder due to required low-level access to hardware. For the same reason steganalysis is harder to perform. Drawback: increased frame error rate.

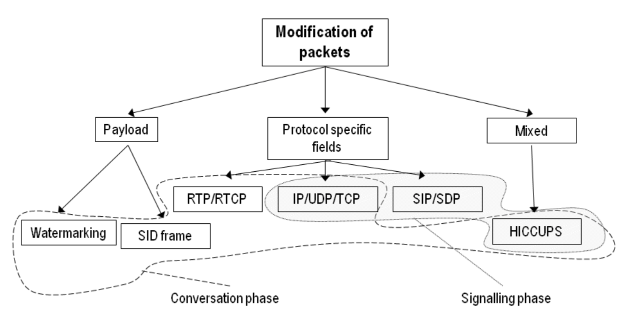

Fig. 5 Classification of steganography based on modification of packets (S1)

Examples of the steganographic methods that modify packets' time dependencies (S2):

- Methods which affect sequence order of packets (in VoIP possible only for RTP).
- Methods which modify inter-packet delay (in VoIP possible for RTP and RTCP; for some protocols, e.g. SIP, not useful due to small number of messages).
- Methods which introduce intentional losses by skipping sequence numbers at sender end (for RTP and RTCP protocols).

Some characteristic features of these methods:

- Sender-receiver synchronization required.
- Lower steganographic capacity and harder to detect than for method which utilize protocol specific fields.
- Straightforward implementation.
- Drawback: potential deterioration of conversation quality.

LACK, described below, is an example of a hybrid steganographic method (S3) which modifies both packets and their time dependencies.

3. STEGANOPHONY

An overview of steganographic methods that may be applied for IP telephony was presented by the authors in papers Mazurczyk and Szczypiorski, 2008a, 2008b.

To summarize, two kinds of solutions are considered:

- New solutions specifically proposed for VoIP, e.g. LACK (Section 3.1).
- Known steganographic methods which were not applied in the VoIP context, e.g.: (1) steganographic methods used for protocols such as SIP/SDP, RTP and RTCP, (2) methods like HICCUPS (Section 3.2), (3) methods which affect time dependencies between VoIP packets.

In the following, a brief overview of three steganographic methods is presented:

- **LACK** which uses intentionally delayed audio packets,
- **HICCUPS** which is a medium-dependent steganographic technique for VoWLAN (Voice over Wireless LAN),
- **VoIP protocols steganography – SIP/SDP** (Session Initiation/Description Protocol) and **RTP/RTCP** (Real-Time Transport/Control Protocol), which utilize the syntax and semantics of protocols.

3.1 LACK

LACK is a hybrid steganographic method since it modifies both packets' content and their time dependencies (Fig. 4).

In general, LACK is intended for a broad class of multimedia, real-time applications, but its main foreseen application (at least for now) is VoIP. The proposed method utilizes the fact that for usual multimedia communication protocols like RTP (Real-Time Transport Protocol) excessively delayed packets are not used for reconstruction of transmitted data at the receiver (the packets are considered useless and discarded). The main idea of LACK is as follows.

At the transmitter, some selected audio packets are intentionally delayed before transmitting. If the delay of such packets at the receiver is considered excessive, the packets are discarded by a receiver not aware of the steganographic procedure. The payload of the intentionally delayed packets is used to transmit secret information to receivers aware of the procedure, so no extra packets are generated. For unaware receivers the hidden data is "invisible".

The idea of LACK is illustrated in Fig. 6. In scenario (1) in Fig. 6, one packet is selected from the RTP stream and its voice payload is substituted with bits of the steganogram. In scenario (2) chosen packets are delayed by a certain value and then sent through the communication channel. In scenario (3), if an excessively delayed packet reaches a receiver unaware of the steganographic procedure, it is discarded. In scenario (4), if the receiver knows about hidden communication, then instead of deleting the packet the receiver extracts the payload.

The effectiveness of LACK depends on many factors such as the details of the communication procedure (in particular the type of codec used, the size of the voice frame, the size of the receiving buffer, etc.) and on the network QoS (packet delay and packet loss probability).

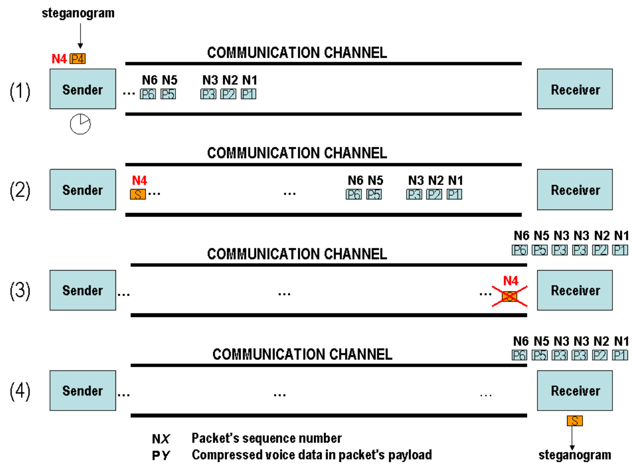

Fig. 6 Idea of LACK

Because legitimate VoIP packets are used it must be realized that conversation quality may be deteriorated. Thus the acceptable level of packet loss for IP telephony must be controlled in the steganographic procedure. The acceptable packet loss is different for various speech codecs, e.g. 1% for G.723.1, 2% for G.729A, 3% for G.711 (if no additional mechanism is used to cope with packet loss). If some additional mechanism to cope with packet loss is used, like PLC (Packet Loss Concealment), then the acceptable loss is higher (e.g. for the G.711 codec it may reach 5-8%).

To be sure that the RTP packet will be recognized as lost at the receiver, the delay must exceed certain value. Two important parameters must be considered and set to the right value: the amount of time by which the chosen packet is delayed (to ensure that it will be considered as lost at the receiver end) and the packet loss probability (to ensure that total packet losses introduced by the network and the LACK procedure will not degrade perceived quality of the conversation). To properly choose the delay value, the capacity of the receiver's de-jitter buffer must be taken into account (the buffer is used to alleviate the jitter effect, i.e. the variations in packets arrival time caused by queuing, contention and serialization in the network). The delay value, usually between 30-70 ms, is important for the end-to-end delay budget, which should not exceed 150 ms.

For example, for the G.711 speech codec with data rate 64 kbit/s and data frame size of 20 ms, if the packet loss probability introduced by LACK is 0.5%, then the practical hidden communication rate is about 320 b/s.

As mentioned above, the performance of LACK depends on the procedure of inserting covert data into the stream of audio packets. The detailed analysis of the dependence of the insertion procedure on the probability distribution of VoIP call duration can be found Mazurczyk and Lubacz, 2008.

LACK, although it is an application layer steganography technique, is less complex to implement than most audio steganography algorithms. The achieved bandwidth is comparable or even higher.

Steganalysis of LACK is harder than in the case of other steganographic methods that were mentioned in this paper. This is mainly because it is common for IP networks to introduce losses. If the amount of packet loss introduced by LACK is kept reasonable, then in principle, it difficult to uncover the hidden communication. Potential steganalysis methods include:

- Statistical analysis of the lost packets for calls in a sub-network. This may be done by a passive warden (or some other network node), based e.g. on RTCP reports (cumulative number of packets lost field) or by observing RTP streams flow (packets' sequence numbers). If for some of the observed calls the number of lost packets is higher than average (or some chosen threshold) this may indicate potential use of LACK.

- An active warden which analyses all RTP streams in the network (SSRC identifier and field: sequence number and timestamp from RTP header) can identify packets that are already too late to be used for voice reconstruction. The active warden may erase their payloads fields or simply drop them. The problem is to avoid eliminating delayed packets that may be used for conversation reconstruction. The size of the jitter buffer at the receiver is not fixed (and may be not constant) and is unknown to the active warden. If an active warden drops all delayed packets, then it can potentially also drop packets that are useful for voice reconstruction, and in effect, the quality of conversation may deteriorate considerably.

3.2 HICCUPS

HICCUPS is a generic steganographic framework for wireless LAN which can be used in voice over wireless LAN (VoWLAN) environments. Information is exchanged in data payloads of frames with intentionally created bad checksums. Normally, stations which do not belong to some hidden group, discard corrupted frames with wrong frame checksums; in HICCUPS these frames carry hidden data and thus enable creating additional on-demand bandwidth for steganographic purposes.

HICCUPS's operation scheme is based on two modes: basic mode and corrupted frame mode. This solution may be utilized in a hidden group which consist of users aware of steganographic procedure. Moreover, the key sequence is common knowledge of the hidden group and is used to switch between two operation modes.

In the basic mode data is exchanged in protocol fields. Such steganographic channel possesses low steganographic capacity – below 1% of available space in the frames. When the key sequence is exchanged, via hidden channels, hidden group stations move from basic to corrupted frame mode. It is worth noting that corrupted frame mode is characterized with much higher steganographic capacity. That is because in the corrupted frame mode information is exchanged in data payload of frames with intentionally created wrong checksums. The method of creating wrong frame checksums is common knowledge for stations from a hidden group of stations. This mode offers almost 100% of available bandwidth for a certain period of time. Normally, stations which do not belong to the hidden group, discard corrupted frames with wrong frame checksums. (Note: some stations may capture all traffic from the network, but a properly adjusted proportion of normal and synthetic distortion, in conjunction with strong cryptography, is sufficient to fool such stations). The next key sequence exchange via hidden channels causes stations from hidden group to return to basic mode.

For a typical case the Frame Error Rate (FER) introduced by the network is about 1.5%. If stations "pretend" that FER is 2.5%, then for a 11 Mbit/s IEEE 802.11b network with 40% usage of bandwidth, the steganographic bandwidth is about 44 kbit/s. For a 54 Mbit/s IEEE 802.11a/g network the steganographic bandwidth is around 216 kbit/s.

3.3 VoIP protocols steganography

This type of steganography covers a wide range of information hiding techniques, including popular techniques based on IP or TCP protocols. The main idea is to use free, redundant or unused fields of these protocols.

The authors have shown (Mazurczyk and Szczypiorski, 2008a, 2008b) how these techniques may be applied to VoIP signalling protocols like SIP/SDP and protocols used in the conversation phase of the call (RTP/RTCP). Such information hiding techniques can be detected by using, for example, active wardens.

An other steganographic technique which may be applied to VoIP protocols is based on utilizing security mechanisms' fields. The main idea is to use authentication tags to transfer data in a covert manner. In the SRTP (Secure RTP, Baugher et al., 2004) standard it is recommended that this field is 80 bits, but smaller values are also acceptable (e.g. 32 bits). A similar method was proposed for IPv6 by Lucena et al., 2005. Altering the content of fields such as authentication tags with steganographic data enables creating covert channels because data in these fields is almost random (due to the cryptographic mechanism operations). This randomness makes it hard to detect hidden information. Only the receiving party, who possesses a pre-shared key (auth_key), is able to detect the hidden data. For overt users (which do not use steganographic methods), wrong authentication data in packets will result in dropping them. Thus most of known steganalysis methods will fail to uncover this type of secret communication. The only solution is to strip off/erase such fields from the packets. But such countermeasures implemented e.g. in active warden causes a serious limitation for overt users. If active warden erases these fields in all packets that it receives then overt users will be unable to provide security services for themselves. Moreover this would be a violation of the active warden rule (that no protocol's semantic or syntax is disrupted).

Consider a scenario for which an authentication tag is 32 bits long and audio packets are generated each 20 ms. For these assumptions, the steganographic capacity is about 1.6 kbit/s; this is a quite high capacity compared to other methods mentioned in this paper.

CONCLUSIONS

No real-world steganographic method is perfect: whatever the method, the hidden information can be potentially discovered. In general, the more hidden information is inserted into the normally transmitted data, the greater the chance it will be detected. But because the number of steganographic methods is large, and there is no single method to detect them, we should consider steganography in VoIP as a threat to public security.

It is thus important to understand the intrinsic nature of various steganographic methods and, in effect, be able to construct effective steganalysis solutions.

The overview presented in this paper may be of some help in this respect. The authors are of the opinion that in the coming years we may expect an intensive growth of the number of new steganographic and steganalysis methods.

ACKNOWLEDGEMENTS

Warsaw University of Technology, Institute of Telecommunications, has collaborated with Dr. Neil John Vallestero, US Army RDECOM CERDEC HQ via contract N62558-07-P-0042.

REFERENCES

Baugher, M., McGrew, D., Naslund, M., Carrara, E., Norrman, K., 2004: The Secure Real-time Transport Protocol (SRTP), IETF, RFC 3711, March 2004

Berk, V., Giani, A., Cybenko, G., 2005: Detection of Covert Channel Encoding in Network Packet Delays, Tech. Rep. TR2005-536, Department of Computer Science, Dartmouth College, November 2005, http://www.ists.dartmouth.edu/library/149.pdf

Chen, B. and Wornell, G. W., 2001: Quantization index modulation: A class of provably good methods for digital watermarking and information embedding, IEEE Trans. Info. Theory, vol. 47, no. 4, pp. 1423–1443, May 2001

Cox, I., Kilian, J., Leighton, F., and Shamoon, T, 1997: Secure spread spectrum watermarking for multimedia. IEEE Transactions on Image Processing 6(12), pp. 1673–1687

Handley, M., Jacobson, V., Perkins, C., 2006: SDP: Session Description Protocol. IETF, RFC 4566, July, 2006

Kundur, D., Ahsan, K., 2003: Practical Internet Steganography: Data Hiding in IP, Proceedings of the Texas Workshop on Security of Information Systems, April 2nd, 2003

Lucena, N., Lewandowski, G., Chapin, S., 2005: Covert Channels in IPv6, In 5th Privacy Enhancing Technologies Workshop, Lecture Notes in Computer Science 3856, pp. 147-166, May 2005

Mazurczyk, W., Lubacz J., 2008: Analysis of a Procedure for Inserting Steganographic Data into VoIP Calls, In Proc. of PGTS '08, the Fifth Polish-German Teletraffic Symposium), Berlin, Germany, October 6-8, 2008

Mazurczyk, W., Szczypiorski, K., 2008a: Covert Channels in SIP for VoIP signalling, In: Hamid Jahankhani, Kenneth Revett, and Dominic Palmer-Brown (Eds.): ICGeS 2008 – Communications in Computer and Information Science (CCIS) 12, Springer Verlag Berlin Heidelberg, Proc. of 4th International Conference on Global E-security 2008, London, United Kingdom, pp. 65-72, June 2008

Mazurczyk, W., Szczypiorski, K., 2008b: Steganography of VoIP Streams. In: R. Meersman and Z. Tari (Eds.): OTM 2008, Part II – Lecture Notes in Computer Science (LNCS) 5332, Springer-Verlag Berlin Heidelberg, Proc. of OnTheMove Federated Conferences and Workshops: The 3rd International Symposium on Information Security (IS'08), Monterrey, Mexico, November 9-14, 2008, pp. 1001-1018

Murdoch, S. J., Lewis, S., 2005: Embedding Covert Channels into TCP/IP. In: Proc. of 7th International Workshop on Information Hiding 2005, LNCS vol. 3727, pp. 247-261, Springer-Verlag, Heidelberg, 2005

Petitcolas, F., Anderson, R., Kuhn, M., 1999: Information Hiding – A Survey, IEEE Special Issue on Protection of Multimedia Content. July 1999

Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A, 2002: SIP: Session Initiation Protocol, IETF, RFC 3261, June 2002

Schulzrinne, H., Casner, S., Frederick, R., Jacobson, V., 2003: RTP: A Transport Protocol for Real-Time Applications, IETF, RFC 3550, July 2003

Servetto, S. D., Vetterli, M. 2001: Communication Using Phantoms: Covert Channels in the Internet, Proc. IEEE International. Symposium on Information Theory (ISIT), June 2001

Szczypiorski, K., 2003: HICCUPS: Hidden Communication System for Corrupted Networks. In Proc. of: ACS'2003, October 22-24, 2003 Międzyzdroje, Poland, pp. 31-40

Szczypiorski, K., Margasinski, I., and Mazurczyk, W., 2007: Steganographic Routing in Multi Agent System Environment, Journal of Information Assurance and Security (JIAS), Dynamic Publishers Inc., Atlanta, GA 30362, USA, Volume 2, Issue 3, September 2007, pp. 235-243, ISSN 1554-1010

Szczypiorski, K., Margasinski, I., Mazurczyk, W., Cabaj, K., Radziszewski, P., 2008: TrustMAS – Trusted Communication Platform for Multi-Agent Systems. In: R. Meersman and Z. Tari (Eds.): OTM 2008, Part II – Lecture Notes in Computer Science (LNCS) 5332, Springer-Verlag Berlin Heidelberg, Proc. of OnTheMove Federated Conferences and Workshops: The 3rd International Symposium on Information Security (IS'08), Monterrey, Mexico, November 9-14, 2008, pp. 1019-1035

Zander, S., Armitage, G., Branch, P., 2007: A Survey of Covert Channels and Countermeasures in Computer Network Protocols. IEEE Communications Surveys & Tutorials, 3rd Quarter 2007, Volume 9, Issue 3, pp. 44-57, ISSN: 1553-87X